

UK Strategy for Data Resources for Social and Economic Research 2009-2012

Veröffentlichungsversion / Published Version

Sonstiges / other

Empfohlene Zitierung / Suggested Citation:

UK Data Forum; Rat für Sozial- und Wirtschaftsdaten (RatSWD). (2009). *UK Strategy for Data Resources for Social and Economic Research 2009-2012*. (RatSWD Working Paper Series, 131). Berlin. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-75352-2>

Nutzungsbedingungen:

Dieser Text wird unter einer CC BY Lizenz (Namensnennung) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier:
<https://creativecommons.org/licenses/by/4.0/deed.de>

Terms of use:

This document is made available under a CC BY Licence (Attribution). For more information see:
<https://creativecommons.org/licenses/by/4.0>

UK Data Forum

Working Paper

UK Strategy for Data Resources for Social and Economic Research 2009-2012

RatSWD Working Paper, No. 131

Provided in Cooperation with:
German Data Forum (RatSWD)

Suggested Citation: UK Data Forum (2009) : UK Strategy for Data Resources for Social and Economic Research 2009-2012, RatSWD Working Paper, No. 131, Rat für Sozial- und Wirtschaftsdaten (RatSWD), Berlin

This Version is available at:
<http://hdl.handle.net/10419/43603>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

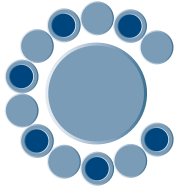
Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



German Council for Social
and Economic Data (RatSWD)

www.ratswd.de

RatSWD

Working Paper Series

Working Paper

No. 131

UK Strategy for Data Resources
for Social and Economic Research
2009-2012

UK Data Forum

December 2009

Working Paper Series of the Council for Social and Economic Data (RatSWD)

The *RatSWD Working Papers* series was launched at the end of 2007. Since 2009, the series has been publishing exclusively conceptual and historical works dealing with the organization of the German statistical infrastructure and research infrastructure in the social, behavioral, and economic sciences. Papers that have appeared in the series deal primarily with the organization of Germany's official statistical system, government agency research, and academic research infrastructure, as well as directly with the work of the RatSWD. Papers addressing the aforementioned topics in other countries as well as supranational aspects are particularly welcome.

RatSWD Working Papers are non-exclusive, which means that there is nothing to prevent you from publishing your work in another venue as well: all papers can and should also appear in professionally, institutionally, and locally specialized journals. The *RatSWD Working Papers* are not available in bookstores but can be ordered online through the RatSWD.

In order to make the series more accessible to readers not fluent in German, the English section of the *RatSWD Working Papers* website presents only those papers published in English, while the German section lists the complete contents of all issues in the series in chronological order.

Starting in 2009, some of the empirical research papers that originally appeared in the *RatSWD Working Papers* series will be published in the series *RatSWD Research Notes*.

The views expressed in the *RatSWD Working Papers* are exclusively the opinions of their authors and not those of the RatSWD.

The RatSWD Working Paper Series is edited by:

Chair of the RatSWD (2007/ 2008 Heike Solga; 2009 Gert G. Wagner)

Managing Director of the RatSWD (Denis Huschka)

Introduction:

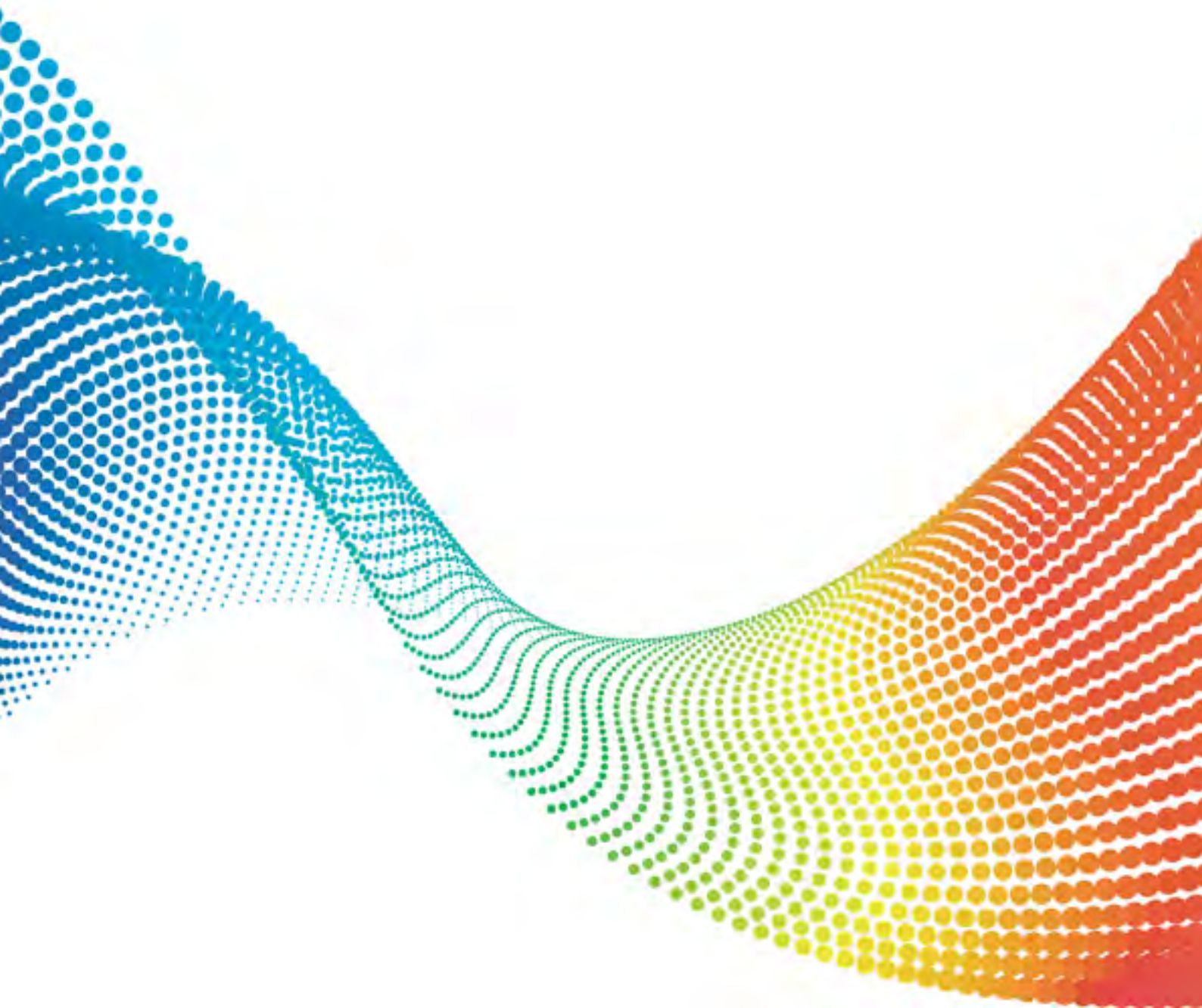
In the United Kingdom a group has been established to help ensure that the right data resources are created and made available to help inform the important choices that decision makers need to make. The group is the UK Data Forum. The members of the Forum come from national research councils, chief statisticians and researchers within government, the national statistics office, national library and national archives, devolved administrations, sub-national bodies and other organisations in the private, public and not for profit sectors.

The current *Strategy for Data Resources for Social and Economic Research: 2009-2012* sets out practical actions on areas for collaboration and the development of cross-cutting facilities. The strategy is owned by the members of the Forum and represents a statement of intent by them.

A central feature of the current strategy, and one of the strongest themes from the consultation undertaken to inform its development, is the internationalisation of the research agenda. Actions in the strategy include initiatives to make collaborative agreements with other agencies within and outside the UK at national, European Union and international levels. The Forum believes that the potential for close working with colleagues in the UK and Germany is likely to be especially valuable. Sharing the strategy with you on this site is a first step in encouraging what we hope will be a deep and fruitful relationship in the future.

John Pullinger

Chair, UK Data Forum



NDS
THE NATIONAL
DATA STRATEGY

UK STRATEGY FOR DATA
RESOURCES FOR SOCIAL
AND ECONOMIC RESEARCH
2009-2012





CONTENTS

01	FOREWORD
03	EXECUTIVE SUMMARY
07	INTRODUCTION
11	SECTION 1 The challenges driving data needs
13	SECTION 2 Building on the foundations of the first National Data Strategy
17	SECTION 3 The digital revolution – proving research value and setting priorities
21	SECTION 4 A strategy for international data requirements
23	SECTION 5 Data access, security, sustainability and ethical considerations
25	SECTION 6 From plans to action: priorities for 2009 to 2012
29	APPENDICES

FOREWORD

IF DECISION-MAKERS IN THE UK ARE TO MAKE SENSIBLE CHOICES ABOUT THE ISSUES THAT FACE OUR COUNTRY THEY NEED A GOOD EVIDENCE BASE. THAT MEANS GOOD DATA. THIS *STRATEGY FOR DATA RESOURCES FOR SOCIAL AND ECONOMIC RESEARCH: 2009-2012* IS A STATEMENT OF INTENT BY ITS NAMED PARTIES TO WORK TOGETHER TO PROVIDE THE DATA THAT WILL BE NECESSARY. IT CONTAINS PRACTICAL ACTIONS ON AREAS FOR COLLABORATION AND THE DEVELOPMENT OF CROSS-CUTTING FACILITIES.

The value to the UK of good data has never been greater. Our consultation highlighted the need to understand instability in economic and financial markets, environmental changes, population dynamics, lifestyles and social trends, to name just some of the headline issues. At a time of increased pressure on resources, the research community must show that investing in it gives a unique rate of return when compared with many competing claims. The case is strong but it must be clear and well-communicated.

The main areas for action we have identified are: the development of longitudinal data; the harnessing of new types of data arising from transactions and tracking records; reconciling the benefits of improved research access to data with the imperative to safeguard confidentiality and privacy; improving geographically referenced data; developing new data on organisations; and, perhaps most challenging of all, support for internationally comparable datasets that allow us to assess the impact on the UK of changes in other countries and to address issues of global importance.

As well as the actions we are taking, we consider that progress in two other areas is essential: user engagement and skills development. In these areas we recognise that a range of other bodies have specific responsibilities. We hope that the actions we are taking will support their activities and look forward to working together with them in the years ahead.

The *Strategy for Data Resources for Social and Economic Research: 2009-2012* is owned by the members of the UK Data Forum and the many contributors to our consultations. I thank them for their active commitment. We owe a particular debt to Peter Elias and Siân Millard who have done all the hard work to support our deliberations and to the Economic and Social Research Council which has continued to fund our endeavour. The first *National Data Strategy*, published in 2007, showed what can be done when the research and policy communities work together. We are well placed now to deliver on a challenging agenda for the years ahead. It matters that we succeed. I look forward to working with all our partners to ensure that we do.

John Pullinger
Chair, UK Data Forum

The UK Strategy for Data Resources for Social and Economic Research 2009-2012 will be taken forward via collaborative engagement between a wide range of research funding bodies, government department and agencies, and the devolved administrations of the United Kingdom. Those currently tasked with this responsibility are:

Andy Sutherland

Head of Profession for Statistics
NHS Information Centre for Health and Social Care

John Pullinger

Librarian and Director General
Information Services, House of Commons

Paul Wiles

Government Chief Social Scientist
Government Social Research Unit, Treasury

Diana Wilkinson

Chief Researcher
The Scottish Government

Keith Dugmore

Director
Demographic Decisions Ltd

Dame Karen Dunnell, DCB

Office for National Statistics

Jude England

Head of Social Science Collections and Research
The British Library

Malcolm Read

JISC Executive Secretary
Joint Information Systems Committee

Norman Caven

Chief Executive
Northern Ireland Statistics and Research Agency

Peter Norris

Head of Data and Analysis
Local Government Association,
Analysis and Research

Richard Bartholomew

Chief Research Officer
Department for Children, Schools and Families

Sharon Jones

Deputy Director, Information Manager
Department for Work and Pensions

Michael Jubb

Director
Research Information Network

Carol Tullo

Director, Information Policy and Services
The National Archives

Sharon Witherspoon MBE

Deputy Director
Nuffield Foundation

Jeremy Neathey

Director for Resources and Training
Economic and Social Research Council

Mark Thorley

Data Management Co-ordinator
Natural Environment Research Council

EXECUTIVE SUMMARY

This document presents the *UK Strategy for Data Resources for Social and Economic Research* (the 'National Data Strategy'). The strategy covers the period 2009-2012, setting out priorities for the development of research data resources both within the social sciences and at the boundaries between the social sciences and other areas of scientific enquiry. It seeks to identify major gaps or weaknesses in the current range of data resources, not just in terms of the research issues that data can address but also the constraints arising from problems of data access, data security and ethical issues surrounding the collection, curation, preservation and/or reuse of particular types of data. It will be monitored and revised towards the end of this period to maintain the well-deserved position the UK has as one of the world's leading nations in terms of the data resources to support social and economic research.

The National Data Strategy is a statement of intent and an action plan, bringing together expressed needs for data from the policy and research communities with current and potential developments in data availability. This planned approach has been undertaken by agencies involved in the funding, development and maintenance of data required for research in the social sciences. The plan is not the property of any one organisation. It represents the collective interests, aims and ambitions of those who have supported its development, including social science researchers in the academic and non-academic sectors, representatives from many government departments, agencies, local government and the devolved administrations, and those research funding bodies with significant interests in the social science research agenda.



The National Data Strategy:

Builds on success – by sustaining major new longitudinal resources and developing improved access to existing and forthcoming cross-sectional data sources

Strengthens recent developments in data services – via measures to support the newly-founded Secure Data Service and Administrative Data Liaison Service and to ensure the continued success of the Economic and Social Data Service

Explores and promotes research use of the new types of data arising from digitisation – specifically transactions data and ‘tracking records’

Encourages the development of procedures, protocols and standards – supporting ethical safeguards surrounding data access and reuse whilst facilitating access for research purposes

Helps to ‘internationalise’ the research agenda – establishing better procedures for data discovery for data held outside the UK and by encouraging use of UK data resources by the international research community

Seeks to improve awareness of the research value of data – about new developments and the potential of existing data resources among research users

Recognises the added complexity and co-ordination requirements resulting from devolution – thereby facilitating comparative research across the countries of the United Kingdom

Plans for the future – by helping to develop a strategic approach to the long term funding, sustainability and preservation of major data resources.

In each of these areas the strategy proposes specific actions to be achieved by a particular date. In some cases steps have already been taken to initiate the cooperation and collaborative working required between the responsible bodies. In other cases further and more detailed investigations are required before these actions can be progressed. In all cases the bodies identified as responsible have agreed to coordinate and take forward the ideas, principles and the preliminary discussions required to move from plan to action.

In date order the actions agreed are as follows:

Action	To be coordinated by:	To be achieved by:
Conduct a review of the need for long term support for cohort studies of young people (see section 6.4)	BIS/ESRC/Devolved administrations	Dec 2009
Set up a Longitudinal Studies Forum to co-ordinate the development of a range of longitudinal studies (see section 6.4)	ESRC/MRC/Wellcome Trust and ONS	March 2010
Establish a long term strategy for the development of a range of longitudinal data resources (see section 6.4)	ESRC/MRC/Wellcome Trust/ONS and other government departments	June 2010
Improve research access to potentially disclosive microdata records on people and organisations whilst protecting confidentiality via the use of new data licensing and access procedures (see sections 6.2 and 6.6)	Government departments/ESRC	June 2010
Develop plans for new data infrastructures for support via the European Union and the European Strategy Forum for Research Infrastructures (see section 6.5)	ESRC/Other European funding agencies	June 2010
Explore the potential research value of new types of data, particularly transactions data and tracking records (see section 6.3)	ESRC/Private sector organisations	Sept 2010
Seek to define, fund and establish a geo-spatial resources advisory service (see section 6.6)	JISC/ESRC	Sept 2010

Action	To be coordinated by:	To be achieved by:
Form collaborative agreements with agencies within and outside the UK, to help establish programmes of secondary data analysis based on data resources held in other countries (see section 6.5)	ESRC/Other funding agencies	September 2010
Develop the International Data Forum – a new body which will seek to promote better access to and sharing of data resources across ‘North/South’ global networks (see section 6.5)	ESRC/Other national funding agencies	June 2011
Revise the National Data Strategy to cover the period 2013-2016 (see section 6.1)	The UKDF	Dec 2011
Develop new information from organisations by conducting a further Workplace Employment Relations Survey and investigate the scope for a new Survey of Management Practices (see section 6.3)	BIS/ESRC	Dec 2011
Establish a Cohort Resources Facility and a new 2012 birth cohort to promote interdisciplinary research across the biomedical and social sciences (see section 6.4)	MRC/Wellcome Trust/ESRC	Oct 2012
Establish new and/or improved access to Census of Population data – both past and future (see section 6.2)	Census Offices/ESRC/JISC	Dec 2012
Review the mechanisms through which knowledge about data and access procedures is communicated to potential users (see section 6.6)	The UKDF	Dec 2011

INTRODUCTION

Research lies at the heart of innovation and knowledge creation. Over the past 20 years, with the growing pace of innovation and the associated transfer of knowledge from research to practical application, research activity in all of the sciences has both broadened and deepened. National and international collaborations between research teams, multidisciplinary approaches to research issues and the development of large-scale research programmes have become commonplace. But it is not just the desire to innovate or to pursue 'knowledge for knowledge's sake' that has stimulated research activity. There is now a growing recognition of the challenges facing human populations which require more intensive research efforts and wider research collaboration.

Across the social, economic and behavioural sciences, the issues that demand more focussed research efforts have become ever more apparent, ranging from the need to understand instability in economic and financial markets and the links between socio-economic behaviour and environmental changes, to population movements in response to poverty, war, disease and natural disasters and the complex relationships between mental and physical health, family development, education and lifestyles.

Research in all of these areas requires data – records and measurements taken from the physical and the human environment. Research data are more than just numbers. They may be quantitative or qualitative, deriving from specialist surveys, experimental techniques, the processing of transactions or from sensing devices.



They may consist of numerical measurements or written text, visual images and audio recordings. Technological advances now provide the opportunity to capture and preserve digital information on a scale which would have been unimaginable just two decades ago. The vast increase in information arising from the digital revolution has the potential to improve further and accelerate research efforts, provided that the requisite data resources needed for scientific research can be collected, marshalled and preserved in ways that facilitate high quality research.

Data resources, whether created specifically for research or policy purposes or generated as the by-product of other activities, are expensive to develop and maintain and often require extensive collaboration between different agencies. The *UK Strategy for Data Resources for Social and Economic Research* (the 'National Data Strategy') addresses these problems by seeking to identify, prioritise and assist with the development of research data in the light of pressing research and policy needs. The strategy helps shape the way in which UK-based organisations with interests in social science¹ research data (data producers, research funding bodies, data archives, etc.) work together and with the research community to maximise the research potential of existing data and to create new resources, developing better access to existing data and facilitating a broad research agenda.

Background

The first National Data Strategy, published in 2007, was developed under the stewardship of the Economic and Social Research Council (ESRC). The second National Data Strategy presented in this document builds upon and extends this earlier work. The process of reviewing the strategy commenced in 2008, informed by workshops with stakeholders and an online consultation with research data users (see Appendix 1). The information so collected has been drawn together by members of the UK Data Forum² – a body which represents the social scientific research data interests of research funding agencies, data producers and users.

¹ Throughout this document the term 'social science' is used to refer to the study of society and the manner in which people behave and impact on the world around us and includes disciplines such as economics, law, sociology, psychology, business studies, education, politics and international studies.

² For details of the Terms of Reference and membership of the UK Data Forum as at August 2009, see Appendix 2.

The need to revise the first Strategy was triggered by a number of important events. Foremost among these is the *Statistics and Registration Service Act 2007*, an important piece of legislation which, by specifying that the UK Office for National Statistics (ONS) should facilitate research and by creating a new legal status for persons requesting research access to microdata held by the ONS, has the potential to promote further the use of national statistics for research purposes. An additional trigger has been the availability of funding through the UK Government's *Large Facilities Capital Fund* – a source of public funding for research infrastructure previously only available to research communities in the physical and biological sciences. Significant earmarked allocations from this source, totalling over £40M in the past three years, are helping to create new data infrastructures for research in the social and biomedical sciences and to promote greater collaboration between the Research Councils, other research funding organisations and government departments. Another factor is the increased use of public sector information for commercial purposes, whereby organisations add value to data in order to produce new information products and services as promoted by the *Re-use of Public Sector Information Regulations 2005*. Finally, a number of recent lapses in data security have caused data owners and providers to examine carefully the procedures they use to preserve the integrity and security of data for which they have responsibilities. This has heightened awareness of the need to ensure that adequate safeguards are in place to protect and preserve the privacy of people and organisations without significantly impeding access to potentially disclosive personal data that are required to inform a growing research agenda.

The plan of this document

The revised Strategy is presented in the following six sections of this document:

Section 1: The challenges driving data needs

Section 2: Building on the foundations of the first National Data Strategy

Section 3: The digital revolution – proving research value and setting priorities

Section 4: A strategy for international data requirements

Section 5: Data access, security, sustainability and ethical considerations

Section 6: From plans to action: priorities for 2009 to 2012

In the interests of brevity, each section outlines the issues, focussing upon the need to resolve specific problems and meet research needs. Readers who wish to pursue issues in more detail or to obtain additional information are directed to the links shown in Appendix 3.



SECTION 1

The challenges driving data needs

Data needs are driven by research and policy interests. Whether for description or more detailed analysis, measurements taken either from the world around us or about the human condition form key parts of the processes of scientific enquiry and policy formulation. The task of determining and prioritising data needs must be linked to the potential research and policy agenda.

This 'research and policy driven' approach guided the development of the first National Data Strategy. Four key research challenges were identified, each of which gave rise to data requirements of various types from a range of sources. These related to: the ageing population; migration; globalisation; and child development.

A number of important data resources that could inform research on these issues were recognised, including: large-scale census and survey data; administrative data relating to health, social security, employment and education; data on organisations; and longitudinal data on individuals, families and households. Longitudinal data, created via successive observations usually over a long period of time, facilitate research on processes of change, their causes and consequences. For all of the listed data resources, the importance of both geographical location of activity and the timeliness of the data were highlighted.

From the online user consultation and the workshops with key data stakeholders (users and producers of data resources) which have informed the revision process, concerns were expressed about the need to adopt a broader and more flexible approach to the identification of future research challenges. In particular, it was argued that the first National Data Strategy did not reflect the growing importance which attaches to the internationalisation of all aspects of the research agenda. Increased concerns about threats to global security, large-scale migration, the global impact of pressures in financial markets and the fluctuating demands for energy, basic materials such as coal, steel and copper from emerging economies, and global health issues all seem likely to form an important part of the international research agenda over the next few years.



Similarly, global efforts to achieve major reductions in green house gas emissions raise issues about trade and competitiveness that will stimulate the need for more and better data on labour and raw material costs, productivity and trade flows. More generally, many stakeholders argued that social scientists must increase their engagement with issues related to environmental changes, whether in response to climate change, pollution, and resource management or in terms of the need to promote and sustain changes in behaviour. Section 4 outlines the steps that will be taken to pursue more vigorously the development of relevant data resources.

Other specific concerns include the need for better access to administrative data at both national and local levels given the potential that exists within such data to inform research on many of the issues listed above, for improved information which assists with the production of population estimates at the detailed spatial level, for better statistics on crime and crime-related behaviour, and for significant improvements in access to high quality data to facilitate research on all aspects of health and wellbeing. In part, these concerns will be met via the new resources and services established under the first National Data Strategy outlined in Section 2.

The issues raised during the consultations with the research and policy community are being fed through to those bodies with responsibilities for the design and development of the UK's major national data resources. In part they will be addressed via strategic planning associated with specific resources. To a large extent, they will also require cooperation and collaboration between various departments and agencies, with support from research funders, to achieve linkage between data sources at the individual, organisational and geographical levels. In turn, this requires that the issues surrounding the ethics of linkage and the security, preservation and sustainability of the linked data are addressed and resolved in ways which maintain public support for these efforts and enhance the research value of the data resources so created. These issues are elaborated further in Section 5.

As well as reacting to expressed needs, the National Data Strategy encourages some proactive developments, particularly via an exploration of the research value of new types of digital data now potentially available for research purposes. Section 3 sets out preliminary plans to encourage innovative research use of new types of digital information.

Section 6 sets out the actions that are required to implement the strategy and identifies the lead organisations that will promote this work. Given the uncertainties facing many organisations in terms of resource availability, these actions may be revised, reordered and reprioritised. Nonetheless, they represent a blueprint for the further improvement of data resources which will inform social and economic research and help guide policy in these and related areas.

SECTION 2

Building on the foundations of the first National Data Strategy

In addition to the expressed needs for new types of data and improved access/linkage arrangements outlined in the preceding section, two important messages arising from our consultation relate to the notions of 'building upon success' and 'making better use of what we have'. These messages inform three dimensions of the strategy: the need to put in place mechanisms to sustain existing resources where these have proved vital for the health of social science research; the need to establish a long-term strategy for the continuing development of large scale data infrastructures that are currently being established; and the need to promote awareness of the wide variety of resources available for potential users in both the research and policy communities.

In terms of 'building upon success', some indication of the importance of the effort that will be required can be gained by considering major new national data sources which are in the process of or have recently been established. Four such developments are highlighted in the boxes opposite:

Understanding Society now replaces the British Household Panel Study. Having started in January 2009, this new longitudinal survey will build up to cover 40,000 households across the UK by 2011, making it the largest household panel study in the world. Funding for the first two rounds of data collection was facilitated via an allocation of £12.5M from the Large Facilities Capital Fund. Significant co-funding for the period 2008/09 to 2011/12 has been offered by a range of government departments in support of the Study.

The **Birth Cohort Facility** is a joint initiative by the ESRC and the Medical Research Council (MRC). The two Councils will be seeking over the coming year to develop the facility in collaboration with other funders, including the Wellcome Trust and government departments. In addition to a new birth cohort being launched around 2012, the facility will enhance the potential for cross-cohort research using earlier birth cohort studies and other longitudinal data resources.

The **Integrated Household Survey** has been developed by the Office for National Statistics (ONS) over the past four years. Starting in January 2009, a core module of questions has been introduced onto the Labour Force Survey, and then gradually extended to cover the National Statistics Omnibus Survey, the General Household Survey – Longitudinal, the Expenditure and Food Survey and the English Housing Survey. The sample size for the Integrated Household Survey will be approximately 200,000 independent households annually, with more than one third of a million interviews conducted with households each year.

The next **Census of Population** will be held on 27 March 2011. Plans are now well advanced and include: the collection of information to identify short-term migrants; new questions on second residences in England and Wales and civil partnerships; retained questions on housing, household composition, employment, occupation, industry and qualifications. When the Census is complete, the Census Longitudinal Study for England and Wales will be updated to provide a longitudinal record covering more than one per cent of individuals in England and Wales located in the 1971, 1981, 1991, 2001 and 2011 censuses. Similar developments, though not covering such a long historical time span, are proceeding in Scotland and Northern Ireland.

In addition to these new or forthcoming data sources, much work has been undertaken over the past few years to develop and enhance the infrastructure that supports the provision of, access to and sustainability of data. This includes continued support for the Economic and Social Data Service and the Census Development Programme, and the launch of two new data services: the Administrative Data Liaison Service and Secure Data Service. Government departments and the Office for National Statistics also play a major role in supporting and enhancing data resources for research purposes. Examples include: *nomisweb* – a service provided by the Office for National Statistics to give free access to detailed and up-to-date UK labour market statistics; *Neighbourhood statistics* – again provided free by the Office for National Statistics and presenting an overview of social, economic, health and environmental indicators for a variety of geographies; the *National Pupil Database*; and the *DWP Longitudinal Study*.

Secure data services – what are they?

Secure data services provide access to sensitive or restricted microdata for research purposes, whilst preventing the release of such data outside the 'safe setting' in which they are held. Traditionally this meant that researchers had to conduct their research within the safe setting established by the data provider. With the advent of secure remote access facilities and techniques to prevent any onward transmission of data, the research user is now able to access such data from a remote location without any breach of confidentiality or release. For examples of such facilities, see:

<http://www.securedata.ukda.ac.uk/about/about.asp> and <http://www.ons.gov.uk/about/who-we-are/our-services/vml>

Given the scale of these developments and the growing importance that will attach to their use as national data resources, a major strategic priority, strongly supported by the many data users and producers who contributed to the process of revising the National Data Strategy, is to ensure that these investments are sustained over the longer-term and that further efforts are undertaken to promote the research value of the data so created and the data services now available. Such efforts should include an understanding that much work remains to be done to improve the existing provision for training of researchers in data management and handling; and where appropriate to provide the incentives or rewards that recognise the work that researchers put into the preparation and sharing of their datasets. This will require concerted and collaborative action by government departments, the Research Councils and other research funding bodies and the ONS. These plans must be coordinated with the research community. Further details indicating how this might be achieved are given in Section 6.



SECTION 3

The digital revolution – proving research value and setting priorities

Technological innovation, particularly the development of digital data capture techniques and the vast increase in data storage capacity at very low cost, has brought about what is often termed the 'data deluge' – the voluminous amounts of information which record such things as the everyday movements of people, the transactions they conduct and the state of the environment they inhabit. The information so generated is not designed for research but addresses the needs of organisations to record, monitor and track activities in the interests of commercial or administrative efficiency and reliability.

This revolution in data capture and storage provides many new possibilities to inform research. But it is not simply the movement away from paper to electronic records that has generated this potential. The rapid spread of the internet and related communications technologies bring with them new types of information that are 'born digital' as the example in the box opposite illustrates.

Example of the research use of new forms of digital data

In the USA, the Centre for Disease Control in Atlanta tracks the progress of influenza outbreaks via data collected and compiled from thousands of health care providers, laboratories and other sources. There is inevitably a time delay to collect and compile such data, so that by the time an upturn is seen, the spread of the disease may be more pronounced.

In 2006 Google labs began analysing trends in terms entered into its search engine. By tracking terms such as 'flu', 'influenza' and 'flu-like symptoms', they have shown not only the close correlation with CDC reports, but have evidence that upturns in trends precedes the CDC report by approximately two weeks, www.google.org/about/flutrends/how.html

Governments, research funding agencies and data regulatory authorities in many countries have been grappling with issues that relate to the research potential arising from the digital information revolution. Some indication of the research value of digital records can be gained by grouping into three broad categories the various types of digital data which have research potential:

Administrative records – a term usually used for data which arise through the administration of a process within the public sector (e.g. a claim for job seeker's allowance, patient health records, student educational progress records, tax records). Such data are generally held by public sector agencies and may be linkable through time at the individual level via personal identifiers (National Insurance numbers, National Health Service numbers, unique student identifiers, etc.).

Commercial records – are generated by organisations usually operating on a 'for-profit' basis and which relate to customer databases, customer usage, and which are generated primarily to improve the efficiency and/or profitability of an organisation. Examples include supermarket loyalty card records, utility usage, bank and financial records.

Tracking records – may be a hybrid of the two above types (e.g. mobile telephone logs) or may be generated by remote sensing devices (e.g. CCTV footage, vehicle movement sensors, web page hits, etc.). Many such data are now generated in 'real-time', often geo-referenced and with potential for linkage at the individual level through time or across geographical space.

Data sharing has become commonplace within the physical sciences (e.g. the Virtual Observatory³ for astronomical data and Genbank⁴ in genomics). These examples are often used to indicate the benefits that can arise from data sharing in some of the physical sciences (e.g. astrophysics, genetics, and climate studies). However, social science data, like medical data, present specific issues that the physical sciences do not usually need to confront. What differentiates such data are their personal nature – the fact that measurements relate to people and/or organisations. This, in turn, creates problems relating to the need to protect identities, the difficulty of gaining access to subjects, the need to respect the wishes and expectations of those subjects, the political sensitivity of the information collected and the associated higher research costs involved in resolving these problems whilst facilitating high quality research.

There is, therefore, a tension between the difficulties of using appropriate methods to capture, process and store data about people and organisations, the problems of prioritising data needs given the sheer volume of information now potentially available for research, the need to preserve security and protect the identities of those persons and organisations to whom the data relate, and the need to ensure that commercial sensitivities and the wishes of individuals are respected.

Administrative data, commercial transactions and ‘tracking’ data all have potential research value, not only because they already exist, but primarily because they cover entire populations. The National Data Strategy recognises the latent research potential in these types of data. Steps will be taken to establish and promote mechanisms to explore their research value whilst overcoming the practical issues associated with reuse. These are outlined in Section 6.

³ See www.worldwidetelescope.org

⁴ See www.ncbi.nlm.nih.gov



SECTION 4

A strategy for international data requirements

One of the strongest themes arising from the consultations with producers and users of data relates to the growing importance which attaches to the internationalisation of the research agenda. There is now a widespread awareness of the need for social scientists, often in collaboration with researchers from other research disciplines, to take a more global approach to many major issues. These include a better understanding of the interdependencies between financial markets and world trade, research on the impacts of attempts to reduce 'green house' gas emissions, the socio-economic consequences of climate change, patterns of migration and their impact on communities, the global spread of infectious diseases, and the international security challenges arising from poverty, war and political instability. Research on these issues requires collaboration across national boundaries involving knowledge exchange, transfer of research expertise and data sharing.



A number of important new developments are already underway to facilitate better access to data and metadata from different countries, further details of which are given in the boxes below:

Council of European Social Science Data Archives (CESSDA) – CESSDA is an umbrella organisation for social science data archives across Europe.

Collectively the constituent CESSDA member organisations serve some 30,000+ social science and humanities researchers and students within the European Research Area each year, providing access to 25,000 data collections, delivering over 70,000 data collections per annum and acquiring a further 1,000 data collections each year. For further information see: <http://www.cessda.org/>

European Social Survey (ESS) – The European Social Survey is an academically driven social survey designed to chart and explain the interaction between Europe's changing institutions and the attitudes, beliefs and behaviour patterns of its diverse populations. Now in its fourth round, the survey covers over 30 nations and employs the most rigorous methodologies.

For further information see: <http://www.cessda.org/>

Significant work has already been undertaken to identify and provide access to a wide range of data resources which can inform a comparative research agenda. While these efforts must continue to be supported, there is now growing concern about the problems that researchers face when seeking to collaborate across national boundaries on issues of global concern. The problems relate to the difficulties of discovering relevant data, of gaining research access to detailed and up-to-date data, of ascertaining the quality of data and of preserving data for reuse. Language barriers often exacerbate these problems despite the growing availability of translation tools.

New initiatives are proposed in Section 6 which will help to build international research data networks to address problems of cross-national data sharing. These include steps to progress international collaboration at the highest levels over issues relating to the sharing of data resources, the resolution of legal problems relating to access to data and the promotion of research networks with the skills and knowledge to make use of the range of data resources potentially available worldwide. Given the need for international co-operation in these areas, these initiatives will require substantial efforts from a number of countries if they are to bear fruit.

SECTION 5

Data access, security, sustainability and ethical considerations

Linkages between different data sources – whether to add medical information to survey data, to enhance the value of census records from individuals or organisations or to delineate more clearly the social, economic or physical environment within which changes in behaviour are observed – are powerful and efficient methods of increasing the research value of data. Key problems to be clarified relate to the *ethics* of linkage and the *security* of the linked data. These interrelated problems have been addressed both by interested parties and by those bodies with responsibilities for maintaining ethical and/or legal safeguards. Whilst the resulting reports and reviews⁵ recognise the research value of linked data, the need for public support for linkage procedures is seen as paramount. There is no simple ‘one size fits all’ solution which will achieve the balance required between the value to society of better research access to more detailed data and improved data linkage procedures on the one hand, and on the other hand the maintenance of data security and ethical safeguards for individual data subjects. Different approaches need to be trialled and evaluated, possibly yielding a variety of mechanisms dependent upon the nature of the data under investigation and the linkages requested.

Work is already underway to address these issues through the use of new access procedures and by establishing a regulated and controlled environment through which access to personal data can be used for *bona fide* research purposes. These are elaborated as follows.

5.1 New developments in data preservation, dissemination, access and licensing arrangements

Following an extensive review of data held by different government departments and as the result of a close collaboration between these departments and the ONS, the ESRC commissioned two new data services in October 2008.

The *Administrative Data Liaison Service* (ADLS) will provide assistance to researchers who wish to make use of or link to administrative data held by government departments and agencies. The *Secure Data Service* (SDS) provides an environment within which researchers can, with the permission of data providers, work on data which may be disclosive of individual or organisational identities while minimising risks of inadvertent disclosure. This service mirrors a similar service currently operated by the ONS (the Virtual Microdata Laboratory).

The development of this ‘safe-setting’ approach to data access can provide reasonably convenient access to data which were previously unavailable for research, or for which access was severely limited. However, the approaches taken by different agencies must be planned, coordinated and appropriately funded if the full benefits of such access arrangements are to be realised. Section 6 describes the actions that will provide the required co-ordination.

In addition to these new services, the UK Data Archive initiated a new self-archiving service (UKDA-store) which enables researchers to archive their research data, metadata and research outputs to promote sharing of data and research results. This service is complementary to the formal preservation and dissemination service for data that is offered by the Economic and Social Data Service (ESDS) for a wide variety of data sources. Through the Secure Data Service, the UKDA also administers a new type of data licensing arrangement which provides ‘ESRC Approved Researcher’ status, with stronger safeguards against the potential misuse of data deemed sensitive or potentially disclosive of the identities of individuals or organisations⁶.

However, some research data can be hard to find and access as not all of the outputs of research, nor material used by researchers, finds its way to the UKDA or HEI repositories. Organisations such as the British Library and The National Archives are working both with and alongside the UKDA and HEIs to support practical initiatives and the establishment of best practice on data management, data security and longer-term data preservation. In a context of rapid technological change, no one solution is likely to provide a definitive answer to the many issues raised by the need for data sustainability and preservation. The UK Data Forum will therefore seek to maintain awareness of and support for initiatives that support data sustainability, access and preservation in the long term.

5.2 Developing data sharing protocols and data access agreements

Data sharing is an efficient and effective way to promote the use of data resources. Data collected for a specific purpose, whether by a government department, the ONS or a separately funded research group may well have much wider research value than that for which it was originally collected. Reusing data via sharing agreements makes enormous sense given the high costs often associated with the original data collection process.

Data sharing agreements need to be administered carefully. The issue here is not just the question of whether data will be misused in some way (e.g. by revealing identities of individual or organisations) and whether there is a lawful basis for data sharing, but the prerogative rights that some researchers feel they have to have the first opportunity to exploit data they have collected and the possible disregard for the legitimate expectations of the subjects. Given the potential benefits for research associated with better sharing and access arrangements, further work will be undertaken to progress these developments whilst safeguarding the data concerned. To this end, the UK Data Forum will initiate a comprehensive review of data access and licensing arrangements, seeking to identify areas where access to data remains problematical and making recommendations for further initiatives where appropriate.

⁵ See, for example, <http://www.justice.gov.uk/reviews/docs/data-sharing-review-report.pdf> and http://www.casimir.org.uk/storyfiles/64.0.data_sharing_code_bamako.pdf

⁶ This licensing procedure should not be confused with licensing for copyright and re-use purposes.

SECTION 6

From plans to action: priorities for 2009 to 2012

The preceding sections have indicated where new or renewed efforts should be placed to carry forward the strategic development of data resources for research across the social sciences. This section elaborates these plans, specifying the actions to be undertaken, the nature of the collaboration required and the indicative timescale over which they should be achieved.

6.1 Sustaining a strategic approach to the development of data resources for research in the social sciences

The consultation exercise indicated that there was considerable support for the wide involvement of researchers, policy analysts, data producers and those responsible for maintaining the 'data infrastructure' in this planned approach to the development of data resources for research across the social sciences. By creating and adopting a more transparent and inclusive strategy, both producers and users of data with research value realise benefits from long-term planning. However, technological changes and research priorities are continually evolving. One of the first priorities therefore is to continue this work, revising the National Data Strategy at appropriate intervals, with the next strategy in place by 2012 and looking forward from 2013 to 2016.

Completion date: December 2011

6.2 Cross-sectional data on individuals and households

The two major sources of information in this area are the responsibility of the ONS in England and Wales. As plans for the Integrated Household Survey reach fruition, the potential for timely and detailed monitoring across a wide variety of aspects of society and the economy will be enhanced. Similarly, plans for the 2011 Census of Population are now approaching completion. This Census will take advantage of numerous developments in data collection and will be the most extensive census ever conducted.

To realise the research potential of these major resources, the following actions are proposed:

- the ONS will work closely with the ESRC to ensure that appropriate and timely methods of data dissemination are available for data from the Integrated Household Survey and from other major national datasets. Where necessary, the ONS will make use of appropriate secure access environments such as those offered by the VML or SDS to provide and approve research access to potentially disclosive data.

Completion date: June 2010

- the three census departments, ESRC, and other relevant funders will co-operate to establish UK-wide access to census outputs (both future and historical) through a variety of ways, including access to linked census records, small area statistics, mobility data, and samples of anonymised records.

Completion date: Ongoing

6.3 Data on and from organisations

Most organisations are legally obliged to provide information to various government departments, agencies and the ONS. This includes financial information for taxation purposes, social security payments, health and safety data and information for regulatory purposes.

In addition, many organisations generate information from their activities and make use of such information to improve business performance. Data on sales, customer characteristics and on purchases provide basic information for this purpose and in some instances will have the potential to inform wider research issues.

In addition to information generated by organisations, there have been some significant attempts by researchers to gather information from organisations by combining traditional survey methods with linkage to administrative and financial records.

To support research on organisations, the following actions will be pursued:

- the Department for Business, Innovation and Skills (BIS) and the ESRC will work together to investigate the scope for a survey looking at management practices and their relation to organisational performance.
Completion date: December 2009
- BIS will also work closely with the ESRC to establish the design of and to commission the next survey in the Workplace Employment Relations series.
Completion date: December 2011

To explore the research value of data collected by organisations, the following steps will be taken:

- the ESRC will investigate the appetite for a collaborative programme of research involving both private and public sector organisations in the use of 'transactions data' – information collected and retained by organisations in the conduct of their business.
Completion date: September 2010

6.4 Longitudinal data on people and households

The success of two bids to the Large Facilities Capital Fund emphasises the importance that attaches to these types of data. Efforts now need to be made to ensure that the scientific and policy communities gain maximum advantage from the increase in resources stemming from new findings and that funding agencies and government departments work together to sustain the investments over time.

To achieve this goal, the following actions will be pursued:

- the MRC, ESRC, Wellcome Trust and the Office for National Statistics will work together to prepare a *Longitudinal Studies Strategy*. This will build upon efforts co-ordinated by the ONS to prepare a strategic view of government interests in such resources and will assist the funding agencies with the preparation of plans for their future funding.
Completion date: March 2010
- the MRC, Wellcome Trust, ESRC in co-operation with the Office for National Statistics and other relevant bodies propose to establish a Longitudinal Studies Forum, to complement the UKDF, which will seek to maintain and enhance the research value of all of the UK's major longitudinal data resources by co-ordinating their future development.
Completion date: June 2010
- ESRC, MRC and the Wellcome Trust, in collaboration with the ONS, will establish a new birth cohort around 2012 and a facility to promote activities that will enhance cross-cohort research generally.
Completion date: October 2012
- in collaboration with BIS and the Department for Children, Schools and Families (DCSF) and the devolved administrations, the ESRC will investigate the feasibility of long-term support for maintenance of cohort studies of young people.
Completion date: December 2009

6.5 Developing international data resources

To promote a more international research agenda the following actions will be undertaken:

- the ESRC will seek to form collaborative agreements with other agencies, both within and outside the UK, to help establish programmes of secondary data analysis based upon data resources in the emerging economies. Further support will also be given to work which seeks to build international networks of researchers, to address issues of global importance, including security, migration, international competitiveness, poverty alleviation, health maintenance and disease control.

Completion date: September 2010

- a group of national research funding agencies will collaborate to establish the need for an *International Data Forum for the Social and Economic Sciences* – an international body designed to facilitate and co-ordinate cross-national production and sharing of data with social science research value by identifying best practice in data sharing and by actively promoting collaborative investments in data collection, new instrumentation, innovative methodologies and globally distributed e-infrastructure.

Completion date: June 2011

- in collaboration with research groups and other research funding agencies in the European Union, the USA and Canada, the ESRC will seek to promote ideas for new research infrastructures which can support and advance the international social science agenda. Where such developments involve a significant number of EU partners, support will be sought for infrastructure development through the European Strategy Forum for Research Infrastructures.

Completion date: June 2010

6.6 Reviewing and promoting existing Data Services and establishing the need for new services

- the ONS and the ESRC will co-operate closely to establish the most efficient use of 'safe-setting' arrangements for access to sensitive and/or disclosive microdata.
- the UK Data Forum (in particular, the Economic and Social Research Council, the Joint Information Systems Committee (JISC), the ONS, The National Archives, the British Library, and the Research Information Network) will review the promotion of existing and newly established data services, seeking to maximise their exploitation and use by the widest community of researchers.

Completion date: December 2011

- the ESRC will work with JISC, DEFRA, the Natural Environment Research Council (NERC) and other relevant bodies to assess the impact of the UK Location Strategy, the work of the UK Location Council and related INSPIRE Directive (Infrastructure for Spatial Information in the European Community). These bodies will explore the feasibility of funding a Geo-spatial Resources Advisory Service – designed to provide a variety of geo-spatial data services to a range of potential research users.

Completion date: September 2010

APPENDICES

.....

APPENDIX 1:
SUMMARY OF CONSULTATION
TO REVIEW THE NATIONAL DATA STRATEGY

.....

APPENDIX 2:
TERMS OF REFERENCE

.....

APPENDIX 3:
ADDITIONAL INFORMATION AND REFERENCE
SOURCES

.....

APPENDIX 4:
LIST OF ACRONYMS

.....

Appendix 1: Summary of Consultation to Review the National Data Strategy

In mid-2008 the ESRC began the process, on behalf of the UKDF, of reviewing the National Data Strategy. The review process proceeded via two strands of engagement: one with data stakeholders such as government departments and devolved administrations through consultation workshops, and the other with users of large-scale data sources through a web-based user survey.

The stakeholder workshops were held in June and October 2008 with the aim of bringing together participants from government departments and agencies and research funding groups to identify, highlight and discuss, from a variety of perspectives, the current and potential future research challenges facing the social sciences. Please see below for a list of participating departments. This evidence gathering was built upon by a User Consultation web survey in July-August 2008 which asked users of large-scale data resources what they considered to be the key social science challenges facing the UK over the coming five years and what data resources would be required to support those challenges. There were over 270 responses to the user consultation.

The UK Data Forum would like to take this opportunity to thank all those who gave their time to feed into this consultation process.

List of participant organisations in the consultation workshops:

British Library
Demographic Decisions Ltd
Department for Children, Families and Schools
Department for Communication and Local Government
Department for Culture, Media and Sport
Department for Work and Pensions
Economic and Social Research Council
Government Office for Science
HM Revenue and Customs
Local Government Association
Medical Research Council
Ministry of Justice
Natural Environment Research Council
NHS Information Centre
Northern Ireland Statistics and Research Agency
Office for National Statistics
Research Information Network
Scottish Executive
The National Archives
Wellcome Trust

Appendix 2: Terms of Reference

The terms of reference of the UK Data Forum are:

- to develop a clear view of the key national and international data resources that are required to meet the future needs of social and economic research and research in related fields of interest
- to identify gaps in the current data infrastructure, and determine how those gaps might be closed
- to prioritise future data provision, including the balance between:
 - enhancing access to and sharing of existing datasets;
 - maintaining ongoing support for major datasets of national importance;
 - new data collection; and
 - withdrawal of support and archiving of existing datasets that have outlived their usefulness.
- to agree principles and strategies for the co-ordinated development and co-funding of prioritised future data resources, as appropriate
- to agree training and capacity building initiatives to promote the effective analysis of data resources
- to identify and seek to resolve the practical and ethical issues associated with linking between data sources.

The UK Data Forum recognises that various concordats between the ESRC and government departments are already in place or are being established and that efforts at data co-ordination are being made across government departments and the devolved bodies. These represent a strong foundation for future collaboration. The Forum builds on and strengthens such work to provide a cross-council, cross-government mechanism for the promotion of collaborative working.

Membership

Chair: Mr John Pullinger Librarian and Director General, Information Services, House of Commons

Member

Mr Richard Bartholomew
 Dr Norman Caven
 Dr Alan Doyle
 Mr Keith Dugmore
 Ms Jude England
 Ms Stephane Goldstein
 Dr Valerie Johnson
 Ms Sharon Jones
 Ms Jil Matheson
 Mr Jeremy Neathey
 Mr Peter Norris
 Mr Malcolm Read
 Dr Andy Sutherland
 Professor Paul Wiles
 Ms Diana Wilkinson
 Dr Sharon Witherspoon

Organisation

Department for Children, Schools and Families
 Northern Ireland Statistics and Research Agency
 Wellcome Trust
 Demographic Decisions Ltd
 British Library
 Research Information Network
 The National Archives
 Department for Work and Pensions
 Office for National Statistics
 ESRC
 Local Government Association
 Joint Information System Committee
 NHS Information Centre
 Home Office
 Scottish Executive
 Nuffield Foundation

Secretariat

Ms Kristine Doronenkova ESRC
 Professor Peter Elias University of Warwick

Appendix 3: Additional information and reference sources⁷

Introduction and background

For details about the development of the first *National Data Strategy* and the *UK Data Forum*, see:

<http://www.esrc.ac.uk/ESRCInfoCentre/NDS/>

To download the first *National Data Strategy*:

http://www.esrc.ac.uk/ESRCInfoCentre/Images/NationalDataStrategyFinal07_tcm6-18160.pdf

To browse the Acts of Parliament referenced in this section, see:

http://www.opsi.gov.uk/acts/acts2007/ukpga_20070018_en_1

<http://www.opsi.gov.uk/advice/psi-regulations/index>

For information about the *Large Facilities Capital Fund* and the *UK Large Facilities Roadmap*:

<http://www.rcuk.ac.uk/research/resinfra/lfcapital.htm>

<http://www.rcuk.ac.uk/cmsweb/downloads/rcuk/publications/lfroadmap08.pdf>

Section 1

The *Economic and Social Research Council* and the *Medical Research Council Strategic Plans for 2009-2014* can be viewed at:

<http://www.esrcsocietytoday.ac.uk/ESRCInfoCentre/strategicplan/>

<http://www.mrc.ac.uk/About/Strategy/StrategicPlan2009-2014/index.htm>

The *Natural Environment Research Council's* strategy for research in the period 2007-2012 can be downloaded at:

<http://www.nerc.ac.uk/publications/strategicplan/documents/strategy07.pdf>

Documents detailing the strategic planning for research and foresight activity for most government departments and the devolved administrations can be found on their websites. Examples from the *Department for Business, Innovation and Skills*, the *Department for Work and Pensions*, the *Department for Children, Schools and Families* and *Scottish Government* are given below:

http://www.dius.gov.uk/partner_organisations/office_for_science/foresight

<http://research.dwp.gov.uk/asd/asd5/#strategy>

<http://publications.dcsf.gov.uk/eOrderingDownload/AES-2009.pdf>

<http://www.scotland.gov.uk/topics/research>

Section 2

For further information about *Understanding Society*, the UK Household Longitudinal Study, see:

<http://www.understandingsociety.org.uk/>

Information about the development of plans for the UK *Birth Cohort Facility* can be found at:

http://www.esrcsocietytoday.ac.uk/ESRCInfoCentre/PO/ESRC_Community/newsarchive/birth_cohort.aspx

The status of the *Integrated Household Survey* is described on the Economic and Social Data Service website at:

<http://www.esds.ac.uk/Government/cps/>

The ESRC's *UK Census of Population Development Programme* is described at:

<http://www.census.ac.uk/> and news is broadcast on <http://twitter.com/Censusacuk>

⁷ All URLs listed in this appendix were checked on 23rd July 2009. If a link appears to be broken, check the Internet Archive 'WayBackMachine' at <http://www.archive.org/index.php> to locate the version of the website archived immediately before this date.

The Office for National Statistics (ONS) maintains *Census of Population* web pages at:
<http://www.ons.gov.uk/census/index.html>

Information about the *Pupil Level Annual Schools Census* and the *National Pupil Database* is given at:
<http://www.bris.ac.uk/Depts/CMPO/PLUG/whatisplug.htm>

Neighbourhood Statistics is an ONS website that provides detailed spatial data across many domains of interest for social scientific research: <http://www.neighbourhood.statistics.gov.uk/dissemination/>

The DWP describes the linked longitudinal database of information gathered from social security and tax records (the *Work and Pensions Longitudinal Study*) at: http://research.dwp.gov.uk/asd/longitudinal_study/ic_longitudinal_study.asp

The work of the *Administrative Data Liaison Service* will be available in the late Autumn 2009 at: <http://www.adls.ac.uk>

The work of the *Secure Data Service* is described at: <http://securedata.ukda.ac.uk/sdata/data.asp>

Nomisweb is a website maintained by the ONS which allows users to download detailed and spatially disaggregated data on the labour market. It can be viewed at: <https://www.nomisweb.co.uk/Default.asp>

Section 3

For details of the work that *Google Labs*® has been undertaking on trends in influenza outbreaks, see:
<http://www.google.org/flutrends/>

Information about an international competition to promote research using existing electronic data is given at:
<http://www.diggingintodata.org/>

Section 4

The *European Roadmap for Research Infrastructures*, developed by the European Strategy Forum on Research Infrastructures (ESFRI), can be downloaded at:
ftp://ftp.cordis.europa.eu/pub/esfri/docs/esfri_roadmap_update_2008.pdf

Information about the *Council of European Social Science Data Archives* is presented at: <http://www.cessda.org/>

Information about the *European Social Survey* can be found at: <http://www.europeansocialsurvey.org/>

Information about the *Survey of Health, Ageing and Retirement in Europe* is at: <http://www.share-project.org/>

Section 5

The Vision Statement from the *UK National Archives* can be downloaded from:
<http://www.nationalarchives.gov.uk/documents/living-information.pdf>

The facility to self-archive data, developed by the *UK Data Archive* is described at: <http://store.data-archive.ac.uk/store/>

The home page for the *UK Economic and Social Data Service* is: <http://www.esds.ac.uk/>

Details about the ONS *secure data laboratory* with remote access can be found at:
<http://www.ons.gov.uk/about/who-we-are/our-services/vml/about-the-vml/vml-overview>

Information about the strategy, policies and programmes of the *British Library* is given at:
<http://www.bl.uk/aboutus/stratpolprog/index.html>

The *Data Sharing Review*, prepared by the Information Commissioner and the Director of the Wellcome Trust is available at: <http://www.justice.gov.uk/reviews/datasharing-intro.htm>

Appendix 4: List of acronyms

ADLS	Administrative Data Liaison Service
BIS	Department for Business, Innovation and Skills
DCSF	Department for Children, Schools and Families
DWP	Department for Work and Pension
ESDS	Economic and Social Data Service
ESRC	Economic and Social Research Council
HEI	Higher Education Institution
INSPIRE	Infrastructure for Spatial Information in the European Community
JISC	Joint Information Systems Committee
MRC	Medical Research Council
NERC	Natural Environment Research Council
ONS	Office for National Statistics
SDS	Secure Data Service
UKDA	United Kingdom Data Archive
UKDF	United Kingdom Data Forum



